Research Division
Federal Reserve Bank of St. Louis
Working Paper Series

Multivariate Contemporaneous Threshold Autoregressive Models

Michael J. Dueker
Zacharias Psaradakis
Martin Sola
and
Fabio Spagnolo

Working Paper 2007-019A
http://research.stlouisfed.org/wp/2007/2007-019.pdf

May 2007

FEDERAL RESERVE BANK OF ST. LOUIS
Research Division
P.O. Box 442
St. Louis, MO 63166

Multivariate Contemporaneous Threshold Autoregressive Models

MICHAEL J. DUEKER

Research Division, Federal Reserve Bank of St. Louis, U.S.A.

ZACHARIAS PSARADAKIS

School of Economics, Mathematics & Statistics, Birkbeck, University of London, U.K.

MARTIN SOLA

School of Economics, Mathematics & Statistics, Birkbeck, University of London, U.K.

Department of Economics, Universidad Torcuato di Tella, Argentina

FABIO SPAGNOLO

Department of Economics and Finance, Brunel University, U.K.

May 2007

Abstract

In this paper we propose a contemporaneous threshold multivariate smooth transition autoregressive (C-MSTAR) model in which the regime weights depend on the ex ante probabilities that latent regime-specific variables exceed certain threshold values. The model is a multivariate generalization of the contemporaneous threshold autoregressive model introduced by Dueker et al. (2007). A key feature of the model is that the transition function depends on all the parameters of the model as well as on the data. The stability and distributional properties of the proposed model are investigated. The C-MSTAR model is also used to examine the relationship between US stock prices and interest rates.

Keywords: Nonlinear autoregressive models; Smooth transition; Stability; Threshold.

JEL Classification: C32; G12.

1 Introduction

Nonlinear time series models which allow for state-dependent or regime-switching behaviour have gained much attention and popularity in recent years. Prominent examples include threshold autoregressive models [see, e.g., Tong (1983)], which are piecewise linear in the threshold space, and Markov switching models [see, e.g., Hamilton (1993)] where regime shifts are driven by a hidden Markov process. Another well-known example is smooth transition autoregressive (STAR) models [see Teräsvirta (1998); van Dijk et al. (2002)] which, unlike threshold or hidden Markov models, allow for smooth rather than discrete changes in regime.[1] More recently, Dueker et al. (2007) introduced a new class of contemporaneous threshold smooth transition autoregressive (C-STAR) models in which the mixing (or regime) weights depend on the ex ante probabilities that regime-specific latent variables exceed certain threshold values. A key feature of the C-STAR model is that its mixing (or transition) function depends on all the parameters of the model as well as on the data, a feature which allows the model to describe time series with a wide variety of conditional distributions.

When the joint dynamic properties of multiple time series are of interest, it is natural to consider multivariate models. In a nonlinear framework, Hamilton (1990), Tsay (1998) and van Dijk et al. (2002), among many others, discussed multivariate Markov switching, threshold and smooth transition autoregressive models, respectively. In spite of some obvious difficulties associated with the practical use of such models (e.g., choice of an appropriate threshold variable, number of regimes, transition function), they are potentially very useful for analyzing possibly state-dependent multivariate relationships. Well-known examples of such relationships, which have been the focus of much research, are nonlinear money-output Granger causality patterns [e.g., Rothman et al. (2001); Psaradakis et al. (2005)] and threshold nonlinearities in the term structure of interest rates [e.g., Tsay (1998); De Gooijer and Vidiella-i-Anguera (2004)], to give but two examples.

[1] Several applications of these models have been proposed in the literature; these include: Tiao and Tsay (1994) and Potter (1995) to US GNP; Rothman (1998), Caner and Hansen (1998) and Koop and Potter (1999) to unemployment rates; Obstfeld and Taylor (1997) to real exchange rates; Aıt-Sahalia (1996), Enders and Granger (1998), Pfann et al. (1996) to interest rates; Pesaran and Potter (1997) to business cycle relationships.

This paper contributes to the literature on multivariate nonlinear models by proposing a contemporaneous threshold multivariate STAR, or C-MSTAR, model. This model is a multivariate generalization of the C-STAR model and shares with the latter the key property that all the variables that are included in the conditioning information set are also present in the mixing function. In analogy with the univariate case, the mixing weights in the C-MSTAR model depend on the ex ante probabilities that latent regime-specific variables exceed certain (unknown) threshold values.

After recalling the definition and main characteristics of univariate C-STAR models in Section 2, the C-MSTAR model is introduced and discussed in Section 3. In particular, we examine the stability properties of the model and give conditions under which the Markov chain associated with the model is geometrically (or, more precisely, Q-geometrically) ergodic. This is a useful property because it implies that the C-MSTAR process is strictly stationary (when suitably initialized) and absolutely regular. We also use artificial data to examine the various types of conditional distributions that can be generated by a C-MSTAR model. In Section 4, we investigate the relationship between US stock prices and interest rates using a C-MSTAR model. Our empirical results suggest that monetary policy has different effects on stock prices in different states of the economy and that Granger causality between stock prices and interest rates is regime dependent. A summary is given in Section 5.

2 Univariate Contemporaneous Threshold Autoregressive Models

The C-STAR model is a member of the STAR family. As is well known, a STAR process may be thought of as a function of two (or more) autoregressive processes which are averaged, at any given point in time, according to a mixing function $G(\cdot)$ with range $[0, 1]$. Specifically, a two-regime (conditionally heteroskedastic) STAR model for the univariate time series $\{x_t\}$ may be formulated as

$$x_t = G(\mathbf{z}_{t-1})x_{1t} + (1 - G(\mathbf{z}_{t-1}))x_{2t}, \quad t = 1, 2, \ldots, \tag{1}$$

where $\mathbf{z}_{t-1}$ is a vector of exogenous and/or pre-determined variables and

$$x_{it} = \mu_i + \sum_{j=1}^{p} \alpha_j^{(i)} x_{t-j} + \sigma_i u_t, \quad i = 1, 2. \tag{2}$$

In (2), $\{u_t\}$ are assumed to be independent and identically distributed (i.i.d.) random variables such that u_t is independent of the past $\{x_{t-1}, x_{t-2}, \ldots\}$ and $\mathbb{E}(u_t) = \mathbb{E}(u_t^2 - 1) = 0$, p is a positive integer, σ_1 and σ_2 are positive constants, and μ_i and $\alpha_j^{(i)}$ $(i = 1, 2; j = 1, \ldots, p)$ are real constants. The feature that differentiates alternative STAR models is the choice of the mixing function $G(\cdot)$ and transition variables $\mathbf{z}_{t-1}$ [cf. Teräsvirta (1998); van Dijk et al. (2002)].

Letting

$$\mathbf{z}_t = (x_t, x_{t-1}, ..., x_{t-p+1})', \qquad \boldsymbol{\delta} = (1, 0, \ldots 0)' \in \mathbb{R}^p,$$

and

$$\mathbf{C}_i = \begin{bmatrix} \alpha_1^{(i)} & \alpha_2^{(i)} & \alpha_3^{(i)} & \cdots & \alpha_{p-1}^{(i)} & \alpha_p^{(i)} \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix}, \quad i = 1, 2,$$

the Gaussian two-regime C-STAR model of order p is obtained by defining the mixing function $G(\cdot)$ in (1) as

$$G(\mathbf{z}_{t-1}) = \frac{\Phi(\{x^* - \mu_1 - \boldsymbol{\delta}'\mathbf{C}_1\mathbf{z}_{t-1}\}/\sigma_1)}{\Phi(\{x^* - \mu_1 - \boldsymbol{\delta}'\mathbf{C}_1\mathbf{z}_{t-1}\}/\sigma_1) + [1 - \Phi(\{x^* - \mu_2 - \boldsymbol{\delta}'\mathbf{C}_2\mathbf{z}_{t-1}\}/\sigma_2)]},$$

where x^* is a threshold parameter and $\Phi(\cdot)$ is the $\mathcal{N}(0,1)$ distribution function.[2] Notice that

$$G(\mathbf{z}_{t-1}) = \frac{\mathbb{P}(x_{1t} < x^* | \mathbf{z}_{t-1}; \boldsymbol{\vartheta}_1)}{\mathbb{P}(x_{1t} < x^* | \mathbf{z}_{t-1}; \boldsymbol{\vartheta}_1) + \mathbb{P}(x_{2t} \geqslant x^* | \mathbf{z}_{t-1}; \boldsymbol{\vartheta}_2)}$$

and

$$1 - G(\mathbf{z}_{t-1}) = \frac{\mathbb{P}(x_{2t} \geq x^* | \mathbf{z}_{t-1}; \boldsymbol{\vartheta}_1)}{\mathbb{P}(x_{1t} < x^* | \mathbf{z}_{t-1}; \boldsymbol{\vartheta}_1) + \mathbb{P}(x_{2t} \geq x^* | \mathbf{z}_{t-1}; \boldsymbol{\vartheta}_2)},$$

where $\boldsymbol{\vartheta}_i = (\mu_i, \alpha_1^{(i)}, \ldots, \alpha_p^{(i)}, \sigma_i^2)'$ is the vector of parameters associated with regime i. Hence, (1) can be rewritten as

$$x_t = \frac{\mathbb{P}(x_{1t} < x^* | \mathbf{z}_{t-1}; \boldsymbol{\vartheta}_1) x_{1t} + \mathbb{P}(x_{2t} \geq x^* | \mathbf{z}_{t-1}; \boldsymbol{\vartheta}_2) x_{2t}}{\mathbb{P}(x_{1t} < x^* | \mathbf{z}_{t-1}; \boldsymbol{\vartheta}_1) + \mathbb{P}(x_{2t} \geq x^* | \mathbf{z}_{t-1}; \boldsymbol{\vartheta}_2)}.$$

[2] Although (conditional) Gaussianity is assumed here and elsewhere in the paper, the Gaussian distribution function could in principle be replaced with another continuous distribution function.

Since the values of the mixing function depend on the probability that the contemporaneous value of x_{1t} (x_{2t}) is smaller (greater) than the threshold level x^*, the model is called a contemporaneous threshold model. As with conventional STAR models, a C-STAR model may be thought of as a regime-switching model that allows for two regimes associated with the two latent variables x_{1t} and x_{2t}. Alternatively, a C-STAR model may be thought of as allowing for a continuum of regimes, each of which is associated with a different value of $G(\mathbf{z}_{t-1})$.

One of the main purposes of the C-STAR model is to address two somewhat arbitrary features of conventional STAR models. First, STAR models specify a delay such that the mixing function for period t consists of a function of x_{t-j} for some $j \geq 1$. Second, STAR models specify which of and in what way the model parameters enter the mixing function. C-STAR models address these twin issues in an intuitive way: they use a forecasting function such that the mixing function depends on the ex ante regime-dependent probabilities that x_t will exceed the threshold value(s). Furthermore, the mixing function makes use of all of the model parameters in a coherent way.

3 Multivariate Contemporaneous Threshold Autoregressive Models

In this section we introduce a multivariate generalization of the C-STAR model. We begin by defining the model and then proceed to investigate some of its properties.

3.1 Definition

The C-MSTAR model proposed in this paper may be viewed as a type of multivariate STAR model. An n-variate (conditionally heteroskedastic) STAR process $\{\mathbf{y}_t\}$ with m regimes may be defined as

$$\mathbf{y}_t = \sum_{i=1}^{m} G_i(\mathbf{z}_{t-1})\mathbf{y}_{it}, \quad t = 1, 2, \ldots, \tag{3}$$

where $G_i(\cdot)$ $(i = 1, \ldots, m)$ are mixing functions with range $[0, 1]$, $\mathbf{z}_{t-1}$ is a vector of exogenous and/or pre-determined variables, and

$$\mathbf{y}_{it} = \boldsymbol{\mu}_i + \sum_{j=1}^{p} \mathbf{A}_j^{(i)} \mathbf{y}_{t-j} + \boldsymbol{\Sigma}_i^{1/2} \mathbf{u}_t, \qquad i = 1, \ldots, m. \tag{4}$$

In (4), $\{\mathbf{u}_t\}$ is a sequence of i.i.d. n-dimensional random vectors such that $\mathbf{u}_t$ is independent of the past $\{\mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \ldots\}$ with $\mathbb{E}(\mathbf{u}_t) = \mathbf{0}$ and $\mathbb{E}(\mathbf{u}_t\mathbf{u}_t') = \mathbf{I}_n$ ($\mathbf{I}_n$ being the n-dimensional identity matrix), p is a positive integer, $\boldsymbol{\mu}_i$ $(i = 1, \ldots, m)$ are n-dimensional vectors of intercepts, $\mathbf{A}_j^{(i)}$ $(i = 1, \ldots, m;\ j = 1, \ldots, p)$ are $n \times n$ coefficient matrices, and $\boldsymbol{\Sigma}_i^{1/2}$ $(i = 1, \ldots, m)$ are symmetric, positive definite $n \times n$ matrices.

For simplicity and clarity of exposition, we shall focus hereafter on the bivariate first-order C-MSTAR model, i.e., the case when $n = 2$, $m = 4$, and $p = 1$. To define this model, let

$$\mathbf{y}_t = (x_t, w_t)', \qquad \mathbf{y}_{it} = (x_{it}, w_{it})' \quad (i = 1, \ldots, 4),$$

$$\mathbf{y}_1^* = (x^*, w^*)', \qquad \mathbf{y}_2^* = (x^*, -w^*)', \qquad \mathbf{y}_3^* = (-x^*, w^*)', \qquad \mathbf{y}_4^* = (-x^*, -w^*)',$$

where x^* and w^* are threshold parameters, and x_{it} and w_{it} $(i = 1, \ldots, 4)$ are latent regime-specific random variables. Then, $\{\mathbf{y}_t\}$ is said to follow a Gaussian first-order C-MSTAR model if it satisfies (3)–(4) with $\mathbf{u}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_2)$, $\mathbf{z}_{t-1} = \mathbf{y}_{t-1}$, and

$$G_i(\mathbf{z}_{t-1}) = (1/\kappa_t)\Phi_2(\boldsymbol{\Sigma}_i^{-1/2}\{\mathbf{y}_i^* - \boldsymbol{\mu}_i - \mathbf{A}_1^{(i)}\mathbf{y}_{t-1}\}), \quad i = 1, \ldots, 4,$$

where $\Phi_2(\cdot)$ is the $\mathcal{N}(\mathbf{0}, \mathbf{I}_2)$ distribution function and

$$\kappa_t = \sum_{i=1}^{4} \Phi_2(\boldsymbol{\Sigma}_i^{-1/2}\{\mathbf{y}_i^* - \boldsymbol{\mu}_i - \mathbf{A}_1^{(i)}\mathbf{y}_{t-1}\}).$$

It can be readily seen that

$$\begin{aligned}
G_1(\mathbf{z}_{t-1}) &= (1/\kappa_t)\mathbb{P}(x_{1t} < x^*, w_{1t} < w^* | \mathbf{y}_{t-1}; \boldsymbol{\theta}_1), \\
G_2(\mathbf{z}_{t-1}) &= (1/\kappa_t)\mathbb{P}(x_{2t} < x^*, w_{2t} \geq w^* | \mathbf{y}_{t-1}; \boldsymbol{\theta}_2), \\
G_3(\mathbf{z}_{t-1}) &= (1/\kappa_t)\mathbb{P}(x_{3t} \geq x^*, w_{3t} < w^* | \mathbf{y}_{t-1}; \boldsymbol{\theta}_3), \\
G_4(\mathbf{z}_{t-1}) &= (1/\kappa_t)\mathbb{P}(x_{4t} \geq x^*, w_{4t} \geq w^* | \mathbf{y}_{t-1}; \boldsymbol{\theta}_4),
\end{aligned}$$

where $\boldsymbol{\theta}_i = (\boldsymbol{\mu}_i', \mathrm{vec}(\mathbf{A}_1^{(i)})', \mathrm{vec}(\boldsymbol{\Sigma}_i)')'$ is the parameter vector associated with regime i. The mixing functions $G_i(\cdot)$ reflect the weighted probabilities that the regime-specific latent variables x_{it} and w_{it} are above or below the respective thresholds x^* and w^*.

3.2 Probabilistic Properties

In this subsection we examine some probabilistic properties of the C-MSTAR model. In particular, we give conditions under which the C-MSTAR model is stable in the sense of having a Markovian representation which is geometrically ergodic.[3] For simplicity and clarity of exposition, we focus once again on the Gaussian, bivariate, first-order C-MSTAR model.

The stability concept employed here is that of Q-geometric ergodicity introduced by Liebscher (2005). To recall the definition of this concept, suppose that $\{\boldsymbol{\xi}_t\}_{t\geq 0}$ is a Markov chain on a general state space $\mathcal{S}$ with k-step transition probability kernel $P^{(k)}(\cdot,\cdot)$ and an invariant distribution $\Pi(\cdot)$, so that $P^{(k)}(\mathbf{x},B) = \mathbb{P}(\boldsymbol{\xi}_k \in B|\boldsymbol{\xi}_0 = \mathbf{x})$ and $\Pi(B) = \int_{\mathcal{S}} P^{(1)}(\mathbf{x},B)\Pi(d\mathbf{x})$ for any Borel set B in $\mathcal{S}$ and $\mathbf{x} \in \mathcal{S}$. Then $\{\boldsymbol{\xi}_t\}$ is said to be Q-geometrically ergodic if there exists a non-negative function $Q(\cdot)$ on $\mathcal{S}$ satisfying $\int_{\mathcal{S}} Q(\mathbf{x})\Pi(d\mathbf{x}) < \infty$ and positive constants a, b and $\lambda < 1$ such that, for all $\mathbf{x} \in \mathcal{S}$,

$$\left\| P^{(k)}(\mathbf{x},\cdot) - \Pi(\cdot) \right\|_{\tau} \leq \{a + bQ(\mathbf{x})\}\lambda^k, \quad k = 1, 2, \ldots,$$

where $\|\cdot\|_{\tau}$ denotes the total variation norm.[4]

Geometric ergodicity entails that the total variation distance between the probability measures $P^{(k)}(\mathbf{x},\cdot)$ and $\Pi(\cdot)$ converges geometrically fast to zero (as k goes to infinity) for all $\mathbf{x} \in \mathcal{S}$. It is well known that, if the initial value $\boldsymbol{\xi}_0$ of the Markov chain has distribution $\Pi(\cdot)$, then geometric ergodicity implies strict stationarity of $\{\boldsymbol{\xi}_t\}$. Furthermore, provided that the initial distribution of $\{\boldsymbol{\xi}_t\}$ is such that $Q(\boldsymbol{\xi}_0)$ is integrable with respect to $\Pi(\cdot)$, Q-geometric ergodicity implies that the Markov chain is Harris ergodic (i.e., aperiodic, irreducible and positive Harris recurrent) as well as absolutely regular (or β-mixing) with a geometrically decaying mixing rate [see Liebscher (2005, Proposition 4)]. Such ergodicity and mixing properties are of much importance for the purposes of statistical inference since they validate the use of well-known asymptotic results [cf. Pötscher and Prucha (1997)].

To give a sufficient condition for Q-geometric ergodicity of a C-MSTAR process, the concept of the joint spectral radius of a set of matrices is needed. Suppose that $\mathcal{C}$ is a

[3] For a comprehensive account of the stability and convergence theory of Markov chains the reader is referred to Meyn and Tweedie (1993).

[4] Note that $\left\| P^{(k)}(\mathbf{x},\cdot) - \Pi(\cdot) \right\|_{\tau} = 2\sup_B \left| P^{(k)}(\mathbf{x},B) - \Pi(B) \right|$.

bounded set of real square matrices and let $\mathcal{C}_h$ be the set of all products of length h ($h \geq 1$) of the elements of $\mathcal{C}$. Then the joint spectral radius of $\mathcal{C}$ is defined as

$$\rho(\mathcal{C}) = \limsup_{h\to\infty} \left(\sup_{\mathbf{C}\in\mathcal{C}_h} \|\mathbf{C}\| \right)^{1/h}, \tag{5}$$

where $\|\cdot\|$ is an arbitrary matrix norm. We note that the value of $\rho(\mathcal{C})$ is independent of the choice of matrix norm and that, if the set $\mathcal{C}$ trivially consists of a single matrix, then $\rho(\mathcal{C})$ coincides with the usual spectral radius (i.e., the largest modulus of the eigenvalues of the matrix).[5]

It is easy to see that the first-order C-MSTAR model

$$\mathbf{y}_t = \sum_{i=1}^{4} G_i(\mathbf{y}_{t-1})(\boldsymbol{\mu}_i + \mathbf{A}_1^{(i)}\mathbf{y}_{t-1}) + \left(\sum_{i=1}^{4} G_i(\mathbf{y}_{t-1})\boldsymbol{\Sigma}_i^{1/2} \right) \mathbf{u}_t, \quad t = 1, 2, \ldots, \tag{6}$$

is a special case of the general nonlinear model considered in Liebscher (2005). Thus, by invoking Theorem 2 of that paper, we have the following result. Here, $\|\cdot\|$ denotes the Euclidean vector norm or the corresponding induced matrix norm (i.e., $\|\mathbf{x}\| = (\mathbf{x}'\mathbf{x})^{1/2}$ and $\|\mathbf{C}\| = \max_{\|\mathbf{x}\|=1} \|\mathbf{C}\mathbf{x}\|$, for any n-dimensional vector $\mathbf{x}$ and $n \times n$ matrix $\mathbf{C}$).

Proposition 1 *Suppose that, for every compact subset B of $\mathbb{R}^2$, there exist positive constants b_1 and b_2 such that $\left\|\boldsymbol{\Sigma}(\mathbf{x})^{-1}\right\| \leq b_1$ and $|\det\{\boldsymbol{\Sigma}(\mathbf{x})\}| \leq b_2$ for all $\mathbf{x} \in B$, where $\boldsymbol{\Sigma}(\mathbf{x}) = \sum_{i=1}^{4} G_i(\mathbf{x})\boldsymbol{\Sigma}_i^{1/2}$. If, in addition, the set $\mathcal{A} = \{\mathbf{A}_1^{(1)}, \mathbf{A}_1^{(2)}, \mathbf{A}_1^{(3)}, \mathbf{A}_1^{(4)}\}$ is such that $\rho(\mathcal{A}) < 1$, then the* C-MSTAR *process $\{\mathbf{y}_t\}$ satisfying* (6) *is a Q-geometrically ergodic Markov chain with $Q(\mathbf{x}) = \|\mathbf{x}\|$.*

It follows from our earlier discussion that $\rho(\mathcal{A}) < 1$ guarantees the existence of a unique invariant distribution for $\{\mathbf{y}_t\}$ with respect to which $\mathbb{E}(\|\mathbf{y}_t\|) < \infty$; furthermore, if $\{\mathbf{y}_t\}$ is initialized from this invariant distribution, then it is strictly stationary as well as absolutely regular at a geometric rate.

It is worth pointing out that Liebscher's (2005) approach, which we have followed here, is quite general and delivers conditions for geometric ergodicity of (conditionally heteroskedastic) nonlinear autoregressive processes which can sometimes be weaker than

[5]By the generalized spectral radius theorem, the matrix norm in the definition of $\rho(\mathcal{C})$ in (5) may be replaced by the spectral radius as long as $\mathcal{C}$ is a finite or bounded set.

alternative sufficient conditions [cf. Liebscher (2005, p. 682)]. A practical difficulty, however, is that exact or approximate computation of the joint spectral radius of a set of matrices is not an easy task, not even in the simplest non-trivial case of a two-element set [see, e.g., Tsitsiklis and Blondel (1997)].[6] One possibility is to use the algorithm presented in Gripenberg (1996) to obtain an arbitrarily small interval within which the joint spectral radius of $\mathcal{A}$ lies. Alternative approximation methods are discussed in Blondel and Nesterov (2005) and Blondel et al. (2005), inter alia.

3.3 Distributional Properties

Some further properties of the C-MSTAR model are illustrated by using the data-generating processes (DGPs) given in Table 1. These DGPs have been chosen to highlight some relevant features of the model with respect to: (i) the response of the mixing function to changes in the parameters of the model; and (ii) the empirical distribution of C-MSTAR data. The errors $\mathbf{u}_t$ are orthogonal under DGP-1, while DGP-2 and DGP-3 allow for positive and negative contemporaneous correlation, respectively. We note that the Q-geometric ergodicity condition of Proposition 1 is satisfied for these DGPs — an application of the algorithm in Gripenberg (1996) yields $0.9366025 < \rho(\mathcal{A}) < 0.9366125$.[7]

Figure 1 shows the conditional density functions of the latent regime-specific random vectors $\mathbf{y}_{it}$ $(i = 1, \ldots, 4)$ for DGP-1, given that $\mathbf{y}_{t-1} = (0.4, 0.6)'$, along with the threshold $\mathbf{y}_1^* = (0.4, 0.6)'$ and the values of the mixing functions $G_i(\mathbf{y}_{t-1})$. Each plot shows the relevant area of the density (suitably rotated) for which each regime is defined. The regime-specific conditional means are $\mathbb{E}(\mathbf{y}_{1t}|\mathbf{y}_{t-1}) = (0.35, 0.57)'$, $\mathbb{E}(\mathbf{y}_{2t}|\mathbf{y}_{t-1}) = (0.29, 0.6)'$, $\mathbb{E}(\mathbf{y}_{3t}|\mathbf{y}_{t-1}) = (0.59, 0.39)'$, and $\mathbb{E}(\mathbf{y}_{4t}|\mathbf{y}_{t-1}) = (0.43, 0.66)'$. It can be seen that the values of the mixing weights $G_i(\mathbf{y}_{t-1})$ depend on the values of the regime-specific conditional means relative to the threshold. More specifically, the larger the area of the conditional distribution which lies above the threshold is, the larger $G_i(\mathbf{y}_{t-1})$ is. In our example, we have $G_1(\mathbf{y}_{t-1}) = 0.09$, $G_2(\mathbf{y}_{t-1}) = 0.48$, $G_3(\mathbf{y}_{t-1}) = 0.09$, and $G_4(\mathbf{y}_{t-1}) = 0.34$.

Conditioning on $\mathbf{y}_{t-1} = (-1.5, -2)'$ results in the density functions shown in Fig-

[6] It should also be remembered that the condition that each of the matrices in $\mathcal{A}$ has a subunit spectral radius is necessary but not sufficient for $\rho(\mathcal{A}) < 1$.

[7] The algorithm is implemented using Gustaf Gripenberg's MATLAB code (which is available at `http://math.tkk.fi/~ggripenb/ggsoftwa.htm`).

ure 2. The regime-specific conditional means are now $\mathbb{E}(\mathbf{y}_{1t}|\mathbf{y}_{t-1}) = (-1.44, -1.97)'$, $\mathbb{E}(\mathbf{y}_{2t}|\mathbf{y}_{t-1}) = (-1.26, -1.97)'$, $\mathbb{E}(\mathbf{y}_{3t}|\mathbf{y}_{t-1}) = (-1.37, -1.35)'$, and $\mathbb{E}(\mathbf{y}_{4t}|\mathbf{y}_{t-1}) = (-1.31, -1.59)'$. The mixing functions take the values $G_1(\mathbf{y}_{t-1}) = 0.88$, $G_2(\mathbf{y}_{t-1}) = 0.1$, $G_3(\mathbf{y}_{t-1}) = 0.02$, and $G_4(\mathbf{y}_{t-1}) = 0$. It is not surprising that the regime associated with $G_1(\cdot)$ is now the most prominent regime since the distance $\mathbb{E}(\mathbf{y}_{1t}|\mathbf{y}_{t-1})$ from each of the thresholds is about one standard deviation.

Figures 3–6 illustrate the effect that contemporaneous correlation has on the mixing functions for the two different conditioning values that were considered before. Notice that, when we condition on $\mathbf{y}_{t-1} = (0.4, 0.6)'$, the values of the mixing functions change substantially as a result of the change in the shape of the conditional distributions. When there is positive correlation $G_1(\mathbf{y}_{t-1}) = 0$, $G_2(\mathbf{y}_{t-1}) = 0.52$, $G_3(\mathbf{y}_{t-1}) = 0.11$, and $G_4(\mathbf{y}_{t-1}) = 0.36$, while $G_1(\mathbf{y}_{t-1}) = 0$, $G_2(\mathbf{y}_{t-1}) = 0.54$, $G_3(\mathbf{y}_{t-1}) = 0.07$, and $G_4(\mathbf{y}_{t-1}) = 0.38$ when there is negative contemporaneous correlation. Interestingly, the change in the sign of the correlation coefficient results in marginal changes in the values of the mixing functions; it is the location of the conditional means relative to the thresholds and the dispersion of the conditional densities that are of primary importance as far as the mixing weights are concerned. Similar results are obtained when we condition on $\mathbf{y}_{t-1} = (-1.5, -2)'$.

3.4 Estimation

As in the univariate case, the parameters of an C-MSTAR model can be estimated by the method of maximum likelihood (ML). For a bivariate first-order model characterized by the parameter vector $\boldsymbol{\theta} = (\boldsymbol{\theta}_1', \boldsymbol{\theta}_2', \boldsymbol{\theta}_3', \boldsymbol{\theta}_4', \mathbf{y}_1^{*\prime})'$, it is not difficult to see that the contribution of the t-th observation to the conditional likelihood is

$$\sum_{i=1}^{4} G_i(\mathbf{y}_{t-1}) \det(\boldsymbol{\Sigma}_i^{-1/2}) \phi_2(\boldsymbol{\Sigma}_i^{-1/2}\{\mathbf{y}_t - \boldsymbol{\mu}_i - \mathbf{A}_1^{(i)}\mathbf{y}_{t-1}\}),$$

where $\phi_2(\cdot)$ is the $\mathcal{N}(\mathbf{0}, \mathbf{I}_2)$ density function. The conditional likelihood function is continuous with respect to the thresholds x^* and w^*, so these parameters can be estimated jointly with all the other parameters of the model. If the C-MSTAR model satisfies the stability condition discussed earlier, so that the data may be assumed to come from a strictly stationary and absolutely regular Markov chain, then it is reasonable to use standard asymptotic procedures to carry out likelihood-based inference on $\boldsymbol{\theta}$.

4 Application: Stock Prices and Interest Rates

As an illustration, we analyze the low-frequency relationship between stock prices and interest rates. The interactions between asset prices and monetary policy is a topic which has attracted considerable interest in the literature [see, e.g., Bernake and Gertler (1999, 2001) and Cecchetti et al. (2000)]. Using a C-MSTAR model, we examine the possibly different effects that monetary policy may have on stock prices in different states of the economy. An interest rate shock, for example, may have very different effects on stock markets depending on whether the price-earnings ratio is (perceived to be) high or low. Our approach explicitly allows for four different regimes, which are associated with: (i) low price-earning ratio, low interest rates; (ii) low price-earning ratio, high interest rates; (iii) high price-earning ratio, low interest rates; and (iv) high price-earning ratio, high interest rates.

More formally, let S_t and R_t denote the ratio of stock prices to earnings per share and the nominal interest rate, respectively. Further, let $s_t = S_t - \mu_s$ and $r_t = R_t - \mu_r$ denote the deviation of the two variables from their respective means. Our analysis is based on the C-MSTAR model

$$\mathbf{y}_t = \sum_{i=1}^{4} G_i(\mathbf{y}_{t-1})\mathbf{y}_{it}, \tag{7}$$

where $\mathbf{y}_t = (s_t, r_t)'$ and $\mathbf{y}_{it} = (s_{it}, r_{it})'$ are latent regime-specific random vectors satisfying

$$\mathbf{y}_{it} = \boldsymbol{\mu}_i + \mathbf{A}_1^{(i)}\mathbf{y}_{t-1} + \boldsymbol{\Sigma}_i^{1/2}\mathbf{u}_t, \quad i = 1, \ldots, 4. \tag{8}$$

In (7)–(8),

$$\begin{aligned}
G_1(\mathbf{y}_{t-1}) &= (1/\kappa_t)\mathbb{P}(s_{1t} < s^*, r_{1t} < r^* | \mathbf{y}_{t-1}; \boldsymbol{\theta}_1), \\
G_2(\mathbf{y}_{t-1}) &= (1/\kappa_t)\mathbb{P}(s_{2t} < s^*, r_{2t} \geq r^* | \mathbf{y}_{t-1}; \boldsymbol{\theta}_2), \\
G_3(\mathbf{y}_{t-1}) &= (1/\kappa_t)\mathbb{P}(s_{3t} \geq s^*, r_{3t} < r^* | \mathbf{y}_{t-1}; \boldsymbol{\theta}_3), \\
G_4(\mathbf{y}_{t-1}) &= (1/\kappa_t)\mathbb{P}(s_{4t} \geq s^*, r_{4t} \geq r^* | \mathbf{y}_{t-1}; \boldsymbol{\theta}_4),
\end{aligned} \tag{9}$$

$$\begin{aligned}
\kappa_t = {} & \mathbb{P}(s_{1t} < s^*, r_{1t} < r^* | \mathbf{y}_{t-1}; \boldsymbol{\theta}_1) + \mathbb{P}(s_{2t} < s^*, r_{2t} \geq r^* | \mathbf{y}_{t-1}; \boldsymbol{\theta}_2) \\
& + \mathbb{P}(s_{3t} \geq s^*, r_{3t} < r^* | \mathbf{y}_{t-1}; \boldsymbol{\theta}_3) + \mathbb{P}(s_{4t} \geq s^*, r_{4t} \geq r^* | \mathbf{y}_{t-1}; \boldsymbol{\theta}_4),
\end{aligned} \tag{10}$$

$$\{\mathbf{u}_t\} \sim \text{i.i.d. } \mathcal{N}(\mathbf{0}, \mathbf{I}_2), \tag{11}$$

and $\boldsymbol{\theta}_i = (\boldsymbol{\mu}_i', \text{vec}(\mathbf{A}_1^{(i)})', \text{vec}(\boldsymbol{\Sigma}_i)')'$.

We use Shiller's (1989) data set of annual observations on the Standard and Poor's 500 composite stock price index to earnings per share (S_t) and the three-month Treasury Bill rate (R_t), extended to cover the period from 1900 to 2000. It is clear from Figure 7 that, for long periods of time, both series take values well above their sample means (which are $\widehat{\mu}_s = 13.731$ and $\widehat{\mu}_r = 4.809$). It is also clear that the series tend to remain above or below the respective sample mean for relatively long periods. It is reasonable to expect that the economy behaved differently in the 1970's and 1980's, when interest rates were relatively high and the price-earnings ratio was relatively low, and in periods such as the 1930's and late 1990's, when the price-earnings ratio was relatively high.

Since we use annual data, we expect that stock price and interest rate dynamics are adequately captured by the first-order model in (7)–(11). ML estimates of the parameters of this model and their asymptotic standard errors (computed from the inverse of the empirical Hessian) are reported in Table 2.[8] The standardized residuals of the model appear to exhibit no signs of serial correlation on the basis of conventional Ljung–Box portmanteau tests.

The estimated threshold parameters reported in the last row of Table 2 are $\widehat{s}^* = 2.0369$ and $\widehat{r}^* = -0.0236$. Adding to these values the corresponding sample means $\widehat{\mu}_s$ and $\widehat{\mu}_r$, we see that the estimated thresholds for the price-earnings ratio and interest rates are 15.7680 and 4.7855, respectively.

The bottom four panels of Figure 7 plots the estimated mixing functions, for each point in sample, which specify the weight of regime 1 (associated with $G_1(\cdot)$), regime 2 (associated with $G_2(\cdot)$), regime 3 (associated with $G_3(\cdot)$), and regime 4 (associated with $G_4(\cdot)$). In Table 4 we date the regimes, attributing a regime to a given time period when the estimated probabilities exceed 0.5 for at least two consecutive observations. It is seen that the most prominent regime is the one characterized by a low price-earnings ratio and low interest rates (regime 1). This regime lasts from mid 1930's to the end of the 1950's. Much of the 1970's and 1980's appear to be associated with a regime with

[8] The ML estimates are obtained by a quasi-Newton optimization algorithm that utilizes the Broyden–Fletcher–Goldfarb–Shano Hessian updating method.

low price-earnings ratio and high interest rates (regime 2), a regime which also seems to characterize a few years in the beginning of the 1920's. The regime associated with high price-earnings ratio and low interest rates (regime 3) never lasts more than six years and is prevalent in only a few years during the 1930's, 1960's and 1990's. Finally, the regime associated with low price-earnings ratio and high interest rates (regime 4) seems to dominate for only a short period of time towards the end of the 1960's, the beginning of the 1970's and the early 1990's.

Regarding the stability properties of the empirical model, we note that the ML estimates reported in Table 2 do not satisfy the condition of Proposition 1; in particular, we have $1.3456313 < \rho(\widehat{\mathcal{A}}) < 1.3765024$, where $\widehat{\mathcal{A}} = \{\widehat{\mathbf{A}}_1^{(1)}, \widehat{\mathbf{A}}_1^{(2)}, \widehat{\mathbf{A}}_1^{(3)}, \widehat{\mathbf{A}}_1^{(4)}\}$. It should be remembered, however, that a subunit joint spectral radius is not necessary for Q-geometric ergodicity.

As an alternative way of assessing the stability of the empirical model, we consider the properties of the noiseless part, or skeleton, of the model [cf. Chan and Tong (1985)]. For the C-MSTAR model in (7)–(11), the skeleton is defined as

$$\overline{\mathbf{y}}_t = F(\overline{\mathbf{y}}_{t-1}, \boldsymbol{\theta}),$$

where

$$F(\overline{\mathbf{y}}_{t-1}, \boldsymbol{\theta}) = \sum_{i=1}^{4} G_i(\overline{\mathbf{y}}_{t-1})(\boldsymbol{\mu}_i + \mathbf{A}_1^{(i)}\overline{\mathbf{y}}_{t-1}).$$

A fixed point of the skeleton is any two-dimensional vector $\overline{\mathbf{y}}_e$ satisfying the equation

$$F(\overline{\mathbf{y}}_e, \boldsymbol{\theta}) = \overline{\mathbf{y}}_e, \tag{12}$$

and $\overline{\mathbf{y}}_e$ is said to be an equilibrium point of the model. Since the model is nonlinear, there may, of course, exist one, several or no equilibrium points satisfying (12). An examination of the local stability of each of the equilibrium points may be carried out by considering the following first-order Taylor expansion around the fixed point:

$$\begin{aligned} \overline{\mathbf{y}}_t - \overline{\mathbf{y}}_e &= F(\overline{\mathbf{y}}_{t-1}, \boldsymbol{\lambda}) - F(\overline{\mathbf{y}}_e, \boldsymbol{\theta}) \\ &\approx \left(\frac{\partial F(\overline{\mathbf{y}}_{t-1}, \boldsymbol{\theta})}{\partial \overline{\mathbf{y}}_{t-1}} \bigg|_{\overline{\mathbf{y}}_{t-1} = \overline{\mathbf{y}}_e} \right)' (\overline{\mathbf{y}}_{t-1} - \overline{\mathbf{y}}_e). \end{aligned} \tag{13}$$

If the matrix of partial derivatives in (13) has a subunit spectral radius, then the equilibrium is locally stable and $\overline{\mathbf{y}}_t$ is a contraction in the neighborhood of $\overline{\mathbf{y}}_e$. It can be readily verified that

$$\frac{\partial F(\overline{\mathbf{y}}_{t-1},\boldsymbol{\theta})}{\partial \overline{\mathbf{y}}_{t-1}} = \sum_{i=1}^{4}\left\{\frac{\partial G_i(\overline{\mathbf{y}}_{t-1})}{\partial \overline{\mathbf{y}}_{t-1}}(\boldsymbol{\mu}_i + \mathbf{A}_1^{(i)}\overline{\mathbf{y}}_{t-1})' + G_i(\overline{\mathbf{y}}_{t-1})(\mathbf{A}_1^{(i)})'\right\} \tag{14}$$

and

$$\frac{\partial G_i(\overline{\mathbf{y}}_{t-1})}{\partial \overline{\mathbf{y}}_{t-1}} = \frac{1}{\kappa_t^2}\left\{-\kappa_t(\boldsymbol{\Sigma}_i^{-1/2}\mathbf{A}_1^{(i)})'\nabla\Phi_2(\mathbf{v}_i) - \Phi_2(\mathbf{v}_i)\sum_{i=1}^{4}(\boldsymbol{\Sigma}_i^{-1/2}\mathbf{A}_1^{(i)})'\nabla\Phi_2(\mathbf{v}_i)\right\}, \tag{15}$$

where $\mathbf{v}_i = \boldsymbol{\Sigma}_i^{-1/2}(\mathbf{y}_i^* - \boldsymbol{\mu}_i - \mathbf{A}_1^{(i)}\overline{\mathbf{y}}_{t-1})$ and $\nabla\Phi_2(\mathbf{v}_i)$ is the gradient of $\Phi_2(\cdot)$ at $\mathbf{v}_i$.

Using numerical simulation and a grid of starting values, it is found that the skeleton of the empirical model in Table 2 has a unique fixed point $\overline{\mathbf{y}}_e = (1.57, 0.05)'$. To assess the stability of the model, we compute the eigenvalues of the matrix of partial derivatives in (13) using the expansion in (14)–(15); these eigenvalues are 0.98 and 0.92, suggesting that the model is locally stable. Furthermore, plots of the skeleton shown in Figure 7 (top panel) reveal that, for both the price-earning ratio and the interest rate, the skeleton converges very quickly to the respective long-run value, thus providing further evidence of stability.

Next, we use the proposed C-MSTAR model to assess the regime-specific Granger causality patterns present in the data. It is important to notice that, using a linear first-order VAR model, the estimated parameters of which are reported in Table 3, none of the two variables appears to be Granger causal for the other. This result is very surprising since, not only do the two variables reflect alternative investing opportunities, but the interest rate is usually thought of as a policy variable that might be used to correct misalignments in stock prices.

Using the C-MSTAR model in Table 2, it can be seen that the elements off the main diagonal of $\mathbf{A}_1^{(i)}$ vary significantly across regimes. Specifically, the interest rate Granger causes the price-earning ratio in regimes 1 and 3 (when the probability of the latent variable r_{it} being below the relevant threshold is high). One may speculate that in regime 3 the stock price boom of the 1960's is associated with a long period of relatively low interest rates; the causality in regime 1 reflects the fact that stocks and bonds are substitute assets. The price-earnings ratio Granger causes the interest rates only in regime 4 (when the probability of r_{4t} and s_{4t} being above their respective thresholds is high). This result may reflect the fact that the central bank reacts to the price-earning ratio by changing the interest rate when it is thought that a misalignment correction is needed. This seems to be captured by our model since regime 4 is usually followed by regime 2. For example, the period of high price-earning ratio and interest rates of the 1920's is followed by a crash in the stock markets.[9]

5 Summary

In this paper we have introduced a new class of contemporaneous threshold multivariate STAR models in which the mixing weights are determined by the probability that contemporaneous latent variables exceed certain threshold values. For a model with first-order dynamics, we have given conditions which ensure that the model is stable in the sense of having a Q-geometrically ergodic Markovian representation. Using numerical examples, we have examined some of the characteristics of the model in terms of the conditional distribution of the data and the properties of the mixing functions. We have also illustrated the practical use of the proposed model by analyzing the bivariate relationship between US stock prices and interest rates.

References

[1] Aït-Sahalia, Y. (1996), Testing Continuous-Time Models of the Spot Interest Rate, *The Review of Financial Studies* 9, 385–426.

[2] Bernanke, B. and Gertler, M. (1999), Monetary policy and asset price volatility, in *New Challenges for Monetary Policy*, Kansas City: Federal Reserve Bank of Kansas City, pp. 77–128.

[3] Bernanke, B. and Gertler, M. (2001), Should central banks respond to movements in asset prices?, *American Economic Review* 91, 253–257.

[9] Even though there is no reason, in general, for regime 4 to be short lived (as this is not an intrinsic property of the model), we expect this to be the case with our data set since a high enough interest rate will tend to cool down the stock markets.

[4] Blondel, V.D. and Nesterov, Y. (2005), Computationally efficient approximations of the joint spectral radius, *SIAM Journal on Matrix Analysis and Applications* 27, 256–272.

[5] Blondel, V.D., Nesterov, Y. and Theys, J. (2005), On the accuracy of the ellipsoid norm approximation of the joint spectral radius, *Linear Algebra and its Applications* 394, 91–107.

[6] Caner, M., and Hansen, B. (1998), Threshold Autoregressions with a Unit Root, *Econometrica* 69, 1555–1597.

[7] Cecchetti S.G., Genberg, H., Lipsky, J. and Wadhwani, S.B. (2000), *Asset Prices and Central Bank Policy*, Geneva Reports on the World Economy, No. 2, International Center for Monetary and Banking Studies and Centre for Economic Policy Research.

[8] Chan, K.S. and Tong, H. (1985), On the use of the deterministic Lyapunov function for the ergodicity of stochastic difference equations, *Advances in Applied Probability* 17, 666–678.

[9] De Gooijer, J.G. and Vidiella-i-Anguera, A. (2004), Forecasting threshold cointegrated systems, *International Journal of Forecasting* 20, 237–253.

[10] Dueker, M. J., Sola, M. and Spagnolo, F. (2007), Contemporaneous threshold autoregressive models: estimation, testing and forecasting, *Journal of Econometrics*, forthcoming.

[11] Enders, W., and Granger, C.W.J. (1998), Unit Root Tests and Asymmetric Adjustment with an Example Using the Term Structure of Interest Rates, *Journal of Business and Economic Statistics* 16, 304–311.

[12] Gripenberg, G. (1996), Computing the joint spectral radius, *Linear Algebra and its Applications* 234, 43–60.

[13] Hamilton, J.D. (1990), Analysis of time series subject to changes in regime, *Journal of Econometrics* 45, 39–70.

[14] Hamilton, J.D. (1993), Estimation, inference and forecasting of time series subject to changes in regime, in Maddala, G.S., Rao, C.R. and H.D. Vinod (eds.), *Handbook of Statistics,* Vol. 11, Amsterdam: Elsevier Science Publishers, pp. 231–260.

[15] Koop, G., and Potter, S.M. (1999), Dynamic Asymmetries in U.S. Unemployment, *Journal of Business and Economic Statistics* 17, 298–312

[16] Liebscher, E. (2005), Towards a unified approach for proving geometric ergodicity and mixing properties of nonlinear autoregressive processes, *Journal of Time Series Analysis* 26, 669–689.

[17] Meyn, S.P. and Tweedie, R.L. (1993), *Markov Chains and Stochastic Stability*, London: Springer-Verlag.

[18] Obstfeld, M., and Taylor, A.M. (1997), Nonlinear Aspects of Goods-Market Arbitrage and Adjustment: Heckscher's Commodity Points Revisited, *Journal of the Japanese and International Economies* 11, 441–479.

[19] Pesaran, M.H., and Potter, S.M. (1997), A Floor and Ceiling Model of U.S. Output, *Journal of Economic Dynamics and Control* 21, 661–695.

[20] Pfann, G.A., Schotman, P.C., and Tschernig, R. (1996), Nonlinear Interest Rate Dynamics and Implications for the Term Structure, *Journal of Econometrics* 74, 149–176.

[21] Pötscher, B.M. and Prucha, I.R. (1997), *Dynamic Nonlinear Econometric Models: Asymptotic Theory*, Berlin: Springer.

[22] Potter, S.M. (1995), A Nonlinear Approach to US GNP, *Journal of Applied Econometrics* 10, 109–125.

[23] Psaradakis, Z., Ravn, M.O. and Sola, M. (2005), Markov switching causality and the money–output relationship, *Journal of Applied Econometrics* 20, 665–683.

[24] Rothman, P. (1998), Forecasting Asymmetric Unemployment Rates, *Review of Economics and Statistics* 80, 164–168.

[25] Rothman, P., van Dijk, D. and Franses, P.H. (2001), A multivariate STAR analysis of the relationship between money and output, *Macroeconomic Dynamics* 5, 506–532.

[26] Shiller, R.J. (1989), *Market Volatility,* Cambridge, Mass.: MIT Press.

[27] Teräsvirta, T. (1998), Modelling economic relationships with smooth transition regressions, in Ullah, A. and D.E.A. Giles (eds.), *Handbook of Applied Economic Statistics*, New York: Marcel Dekker, pp. 507–552.

[28] Tiao, G. C., and Tsay, R. S. (1994), Some Advances in Non-Linear and Adaptive Modelling in Time-Series, *Journal of Forecasting* 13, 109–131.

[29] Tong, H. (1983), *Threshold Models in Non-Linear Time Series Analysis*, New York: Springer-Verlag.

[30] Tsay, R.S. (1998), Testing and modeling multivariate threshold models, *Journal of the American Statistical Association* 93, 1188–1202.

[31] Tsitsiklis, J.N. and Blondel, V.D. (1997), The Lyapunov exponent and joint spectral radius of pairs of matrices are hard – when not impossible – to compute and to approximate, *Mathematics of Control, Signals, and Systems* 10, 31–40.

[32] van Dijk, D., Teräsvirta, T. and Franses, P.H. (2002), Smooth transition autoregressive models - a survey of recent developments, *Econometric Reviews* 21, 1–47.

Table 1. Data-Generating Processes

DGP-1

$$\boldsymbol{\mu}_1 = \begin{bmatrix} -0.05 \\ -0.05 \end{bmatrix}, \quad \mathbf{A}_1^{(1)} = \begin{bmatrix} 0.80 & 0.05 \\ 0.10 & 0.90 \end{bmatrix}, \quad \boldsymbol{\Sigma}_1 = \mathbf{I}_2$$

$$\boldsymbol{\mu}_2 = \begin{bmatrix} -0.05 \\ 0.05 \end{bmatrix}, \quad \mathbf{A}_1^{(2)} = \begin{bmatrix} 0.75 & -0.05 \\ 0.05 & 0.85 \end{bmatrix}, \quad \boldsymbol{\Sigma}^{(2)} = \mathbf{I}_2$$

$$\boldsymbol{\mu}_3 = \begin{bmatrix} 0.15 \\ -0.05 \end{bmatrix}, \quad \mathbf{A}_1^{(3)} = \begin{bmatrix} 0.75 & -0.30 \\ 0.20 & 0.85 \end{bmatrix}, \quad \boldsymbol{\Sigma}^{(3)} = \mathbf{I}_2$$

$$\boldsymbol{\mu}_4 = \begin{bmatrix} 0.05 \\ 0.10 \end{bmatrix}, \quad \mathbf{A}_1^{(4)} = \begin{bmatrix} 0.90 & -0.10 \\ 0.01 & 0.90 \end{bmatrix}, \quad \boldsymbol{\Sigma}^{(4)} = \mathbf{I}_2$$

$$(x^*, w^*) = (0.6, -0.4)$$

DGP-2

Intercepts, autoregressive coefficients and threshold parameters are the same as for DGP-1.

$$\boldsymbol{\Sigma}_1 = \begin{bmatrix} 1 & 0.9 \\ 0.9 & 1 \end{bmatrix}, \ \boldsymbol{\Sigma}_2 = \begin{bmatrix} 1 & 0.8 \\ 0.8 & 1 \end{bmatrix}, \ \boldsymbol{\Sigma}_3 = \begin{bmatrix} 1 & 0.3 \\ 0.3 & 1 \end{bmatrix}$$
$$\boldsymbol{\Sigma}_4 = \begin{bmatrix} 1 & 0.8 \\ 0.8 & 1 \end{bmatrix}$$

DGP-3

Intercepts, autoregressive coefficients and threshold parameters are the same as for DGP-1.

$$\boldsymbol{\Sigma}_1 = \begin{bmatrix} 1 & -0.9 \\ -0.9 & 1 \end{bmatrix}, \ \boldsymbol{\Sigma}_2 = \begin{bmatrix} 1 & -0.8 \\ -0.8 & 1 \end{bmatrix}, \ \boldsymbol{\Sigma}_3 = \begin{bmatrix} 1 & -0.3 \\ -0.3 & 1 \end{bmatrix}$$
$$\boldsymbol{\Sigma}_4 = \begin{bmatrix} 1 & -0.8 \\ -0.8 & 1 \end{bmatrix}$$

Table 2. ML Estimates for a C-MSTAR Model

Regime 1: Low Price-Earning Ratio, Low Interest Rate

$$\widehat{\boldsymbol{\mu}}_1 = \begin{bmatrix} \underset{(0.7752)}{-1.0289} \\ \underset{(0.0769)}{0.4723} \end{bmatrix}, \quad \widehat{\mathbf{A}}_1^{(1)} = \begin{bmatrix} \underset{(0.1289)}{1.1534} & \underset{(0.2525)}{-0.5527} \\ \underset{(0.0125)}{0.0198} & \underset{(0.0241)}{1.0912} \end{bmatrix}, \quad \widehat{\boldsymbol{\Sigma}}_1 = \begin{bmatrix} \underset{(0.7658)}{2.5784} & \underset{(0.0121)}{-0.0079} \\ \underset{(0.0121)}{-0.0079} & \underset{(0.0431)}{0.0226} \end{bmatrix}$$

Regime 2: Low Price-Earning Ratio, High Interest Rate

$$\widehat{\boldsymbol{\mu}}_2 = \begin{bmatrix} \underset{(0.4020)}{1.5035} \\ \underset{(0.4496)}{0.1419} \end{bmatrix}, \quad \widehat{\mathbf{A}}_1^{(2)} = \begin{bmatrix} \underset{(0.1014)}{1.0958} & \underset{(0.0658)}{-0.0210} \\ \underset{(0.1338)}{0.0039} & \underset{(0.0554)}{0.6712} \end{bmatrix}, \quad \widehat{\boldsymbol{\Sigma}}_2 = \begin{bmatrix} \underset{(0.3983)}{1.5199} & \underset{(0.3498)}{0.5858} \\ \underset{(0.3498)}{0.5858} & \underset{(0.1813)}{1.2808} \end{bmatrix}$$

Regime 3: High Price-Earning Ratio, Low Interest Rate

$$\widehat{\boldsymbol{\mu}}_3 = \begin{bmatrix} \underset{(1.6322)}{-1.1839} \\ \underset{(0.5170)}{-0.6837} \end{bmatrix}, \quad \widehat{\mathbf{A}}_1^{(3)} = \begin{bmatrix} \underset{(0.2126)}{1.1327} & \underset{(0.3966)}{1.5213} \\ \underset{(0.0658)}{0.0604} & \underset{(0.1243)}{0.9141} \end{bmatrix}, \quad \widehat{\boldsymbol{\Sigma}}_3 = \begin{bmatrix} \underset{(3.0768)}{8.6217} & \underset{(0.3365)}{0.1760} \\ \underset{(0.3365)}{0.1760} & \underset{(0.5218)}{0.7830} \end{bmatrix}$$

Regime 4: High Price-Earning Ratio, High Interest Rate

$$\widehat{\boldsymbol{\mu}}_4 = \begin{bmatrix} \underset{(1.2241)}{1.0271} \\ \underset{(0.5970)}{1.1669} \end{bmatrix}, \quad \widehat{\mathbf{A}}_1^{(4)} = \begin{bmatrix} \underset{(0.2287)}{0.4349} & \underset{(0.3719)}{-0.0996} \\ \underset{(0.1034)}{-0.3988} & \underset{(0.1841)}{0.8856} \end{bmatrix}, \quad \widehat{\boldsymbol{\Sigma}}_4 = \begin{bmatrix} \underset{(7.0952)}{22.3045} & \underset{(1.4534)}{-3.4039} \\ \underset{(1.4534)}{-3.4039} & \underset{(1.2573)}{3.2939} \end{bmatrix}$$

$$\widehat{s}^* = \underset{(0.5487)}{2.0369}, \qquad \widehat{r}^* = \underset{(0.2027)}{-0.0236}, \qquad \max L = -362.941$$

Figures in parentheses are asymptotic standard errors and max$L$ is the maximized log-likelihood.

Table 3. ML Estimates for a VAR Model

$$\mathbf{y_t} = \boldsymbol{\mu} + \mathbf{A}\mathbf{y_{t-1}} + \boldsymbol{\Sigma}^{1/2} u_t$$

$$\widehat{\boldsymbol{\mu}} = \begin{bmatrix} \underset{(0.3200)}{0.1301} \\ \underset{(0.1503)}{0.0111} \end{bmatrix}, \quad \widehat{\mathbf{A}} = \begin{bmatrix} \underset{(0.0706)}{0.7938} & \underset{(0.0332)}{-0.0590} \\ \underset{(0.1047)}{0.0988} & \underset{(0.0492)}{0.8661} \end{bmatrix}, \quad \widehat{\boldsymbol{\Sigma}} = \begin{bmatrix} 10.2291 & 0.0577 \\ 0.0577 & 2.2561 \end{bmatrix}$$

max$L = -437.679$

Figures in parentheses are asymptotic standard errors and max$L$ is the maximized Gaussian log-likelihood.

Table 4. Dating of Regimes

Regime 1	Regime 2	Regime 3	Regime 4
1934–1959	1919–1924	1931–1935	1969–1972
	1975–1990	1961–1967	1991–1992
		1993–1994	
		1999–2000	

Regime 1: Low price-earning ratio, low interest rate.

Regime 2: Low price-earning ratio, high interest rate.

Regime 3: High price-earning ratio, low interest rate.

Regime 4: High price-earning ratio, high interest rate.

DGP1: Distributions Conditional on $x_{t-1}=.4$ and $w_{t-1}=.6$

$G_1(y_{t-1}) = 0.09 \quad G_2(y_{t-1}) = 0.48 \quad G_3(y_{t-1}) = 0.09 \quad G_4(y_{t-1}) = 0.34 \quad x^* = 0.6 \quad w^* = -0.4$

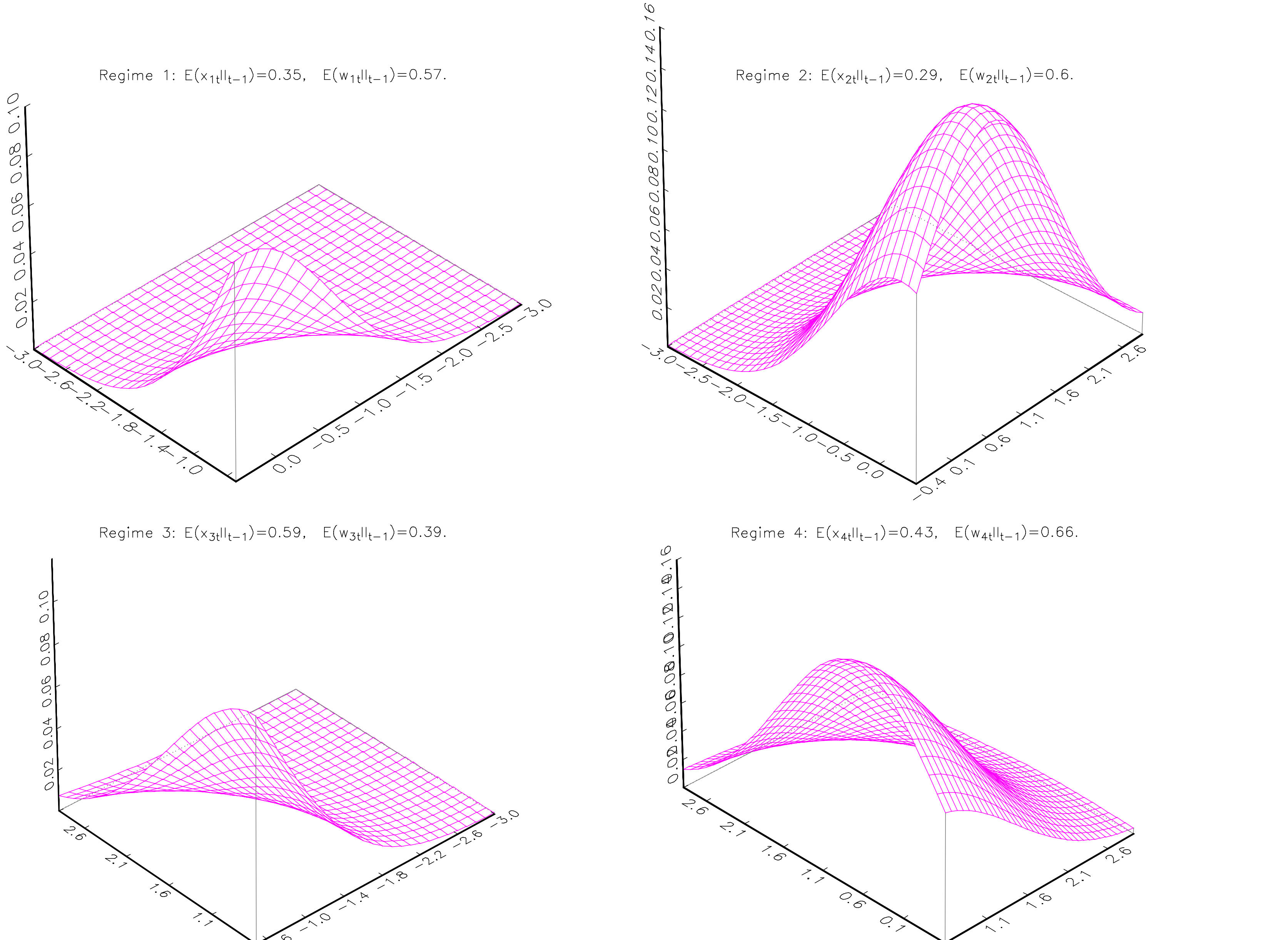

Figure 1

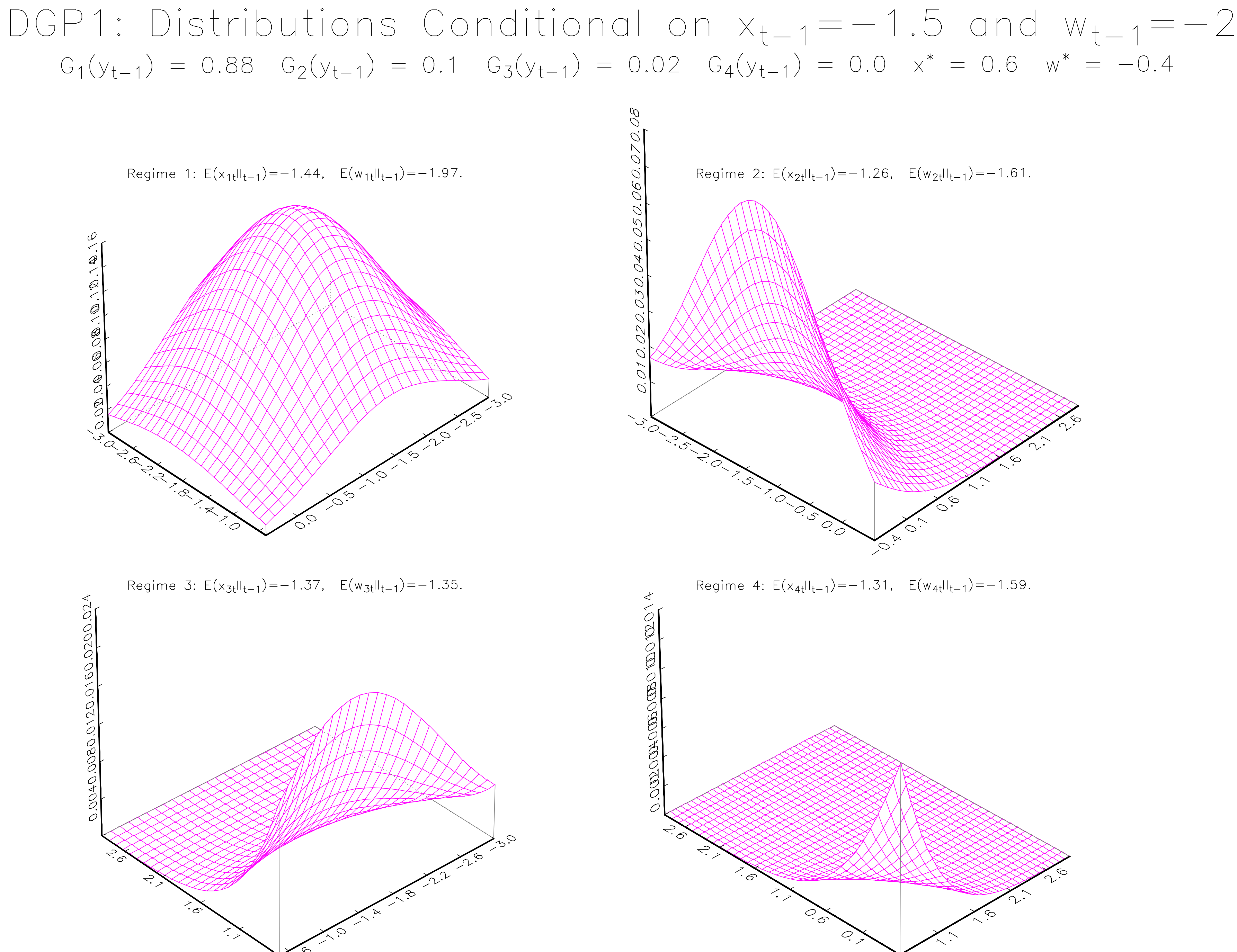
DGP1: Distributions Conditional on $x_{t-1}=-1.5$ and $w_{t-1}=-2$
$G_1(y_{t-1}) = 0.88$ $G_2(y_{t-1}) = 0.1$ $G_3(y_{t-1}) = 0.02$ $G_4(y_{t-1}) = 0.0$ $x^* = 0.6$ $w^* = -0.4$
Regime 1: $E(x_{1t}|I_{t-1})=-1.44$, $E(w_{1t}|I_{t-1})=-1.97$.
Regime 2: $E(x_{2t}|I_{t-1})=-1.26$, $E(w_{2t}|I_{t-1})=-1.61$.
Regime 3: $E(x_{3t}|I_{t-1})=-1.37$, $E(w_{3t}|I_{t-1})=-1.35$.
Regime 4: $E(x_{4t}|I_{t-1})=-1.31$, $E(w_{4t}|I_{t-1})=-1.59$.

Figure 2

DGP2: Distributions Conditional on $x_{t-1}=.4$ and $w_{t-1}=.6$

$G_1(y_{t-1}) = 0.00 \quad G_2(y_{t-1}) = 0.52 \quad G_3(y_{t-1}) = 0.11 \quad G_4(y_{t-1}) = 0.36 \quad x^* = 0.6 \quad w^* = -0.4$

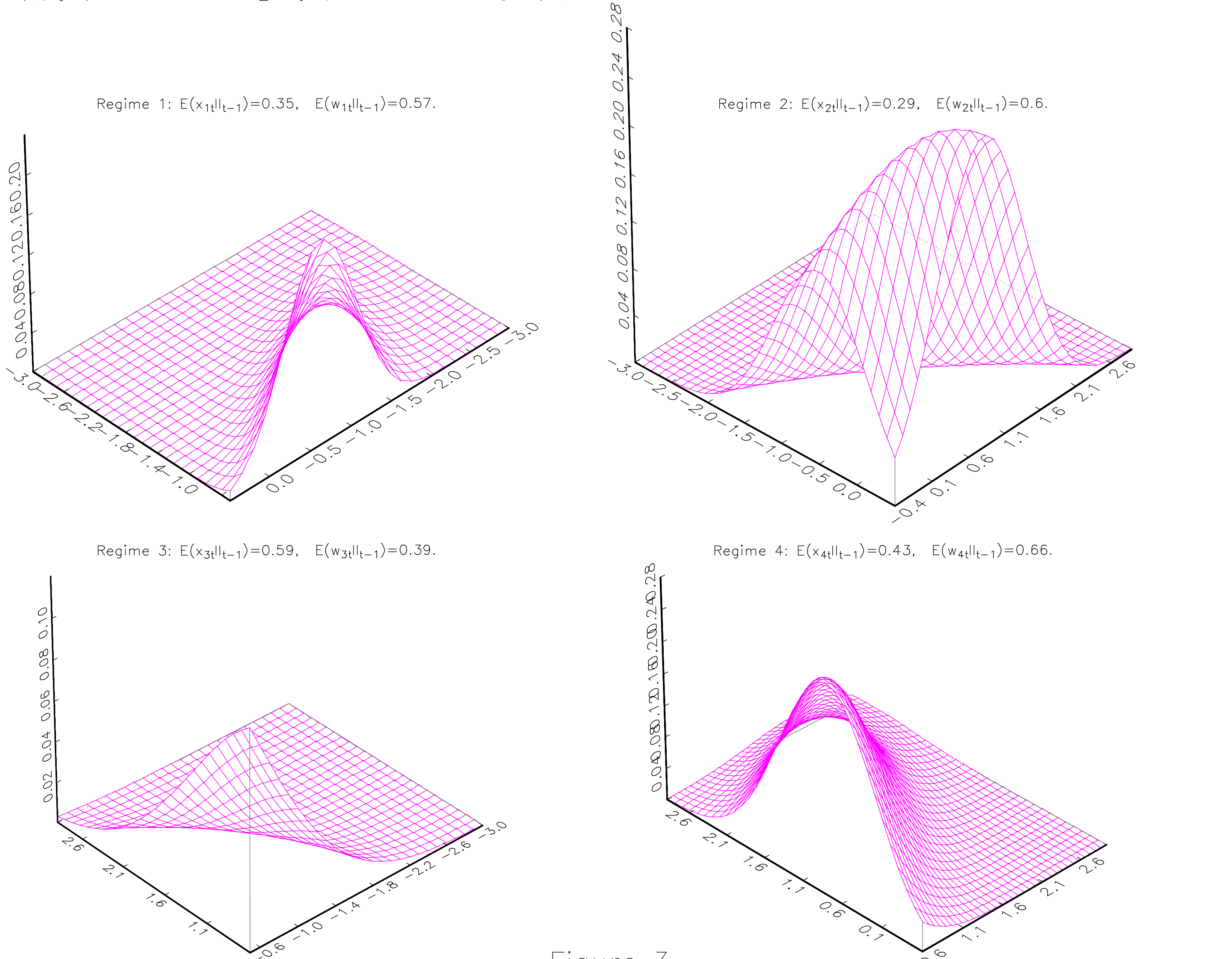

Figure 3

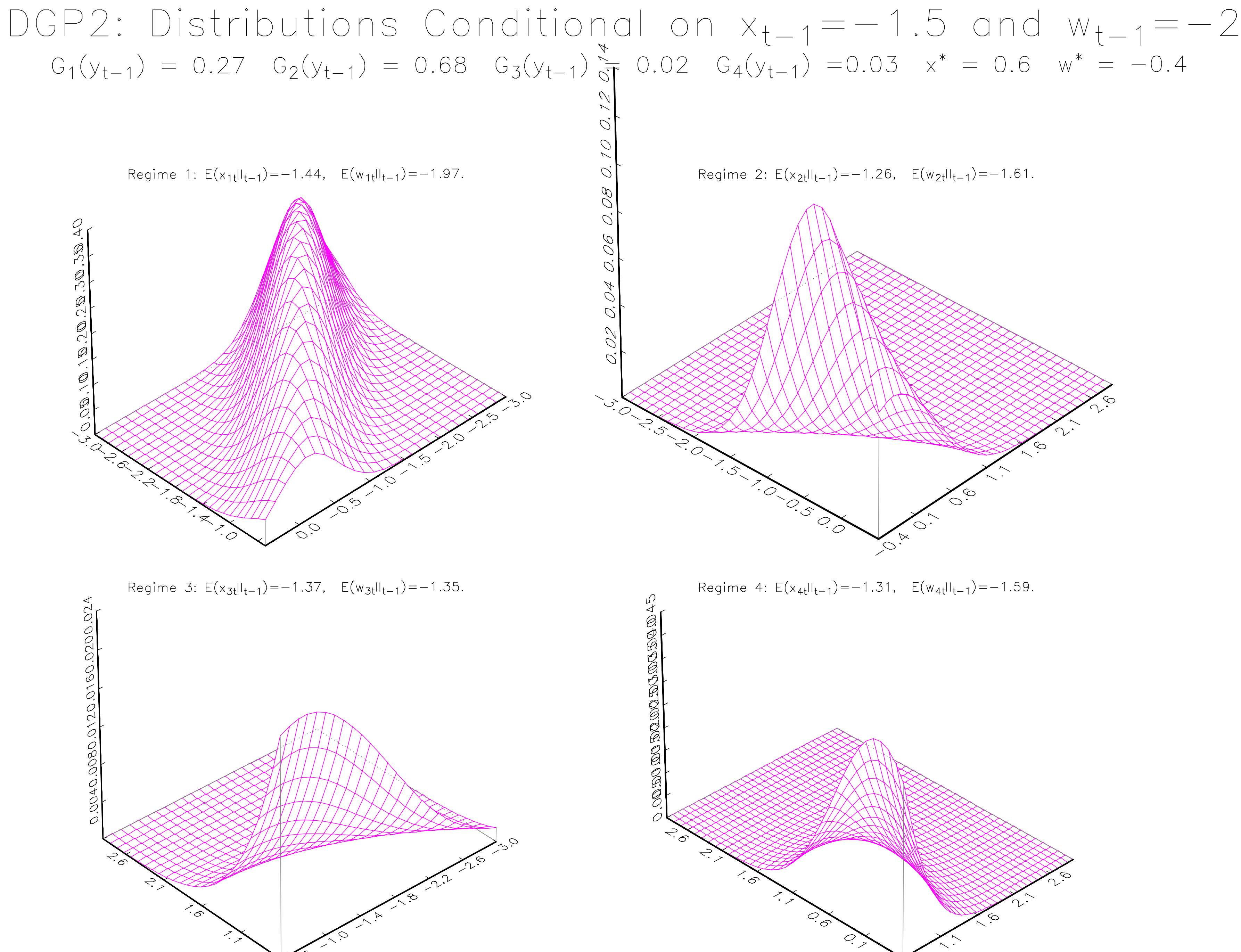
DGP2: Distributions Conditional on $x_{t-1}=-1.5$ and $w_{t-1}=-2$
$G_1(y_{t-1}) = 0.27$ $G_2(y_{t-1}) = 0.68$ $G_3(y_{t-1}) = 0.02$ $G_4(y_{t-1}) = 0.03$ $x^* = 0.6$ $w^* = -0.4$
Regime 1: $E(x_{1t}|I_{t-1})=-1.44$, $E(w_{1t}|I_{t-1})=-1.97$.
Regime 2: $E(x_{2t}|I_{t-1})=-1.26$, $E(w_{2t}|I_{t-1})=-1.61$.
Regime 3: $E(x_{3t}|I_{t-1})=-1.37$, $E(w_{3t}|I_{t-1})=-1.35$.
Regime 4: $E(x_{4t}|I_{t-1})=-1.31$, $E(w_{4t}|I_{t-1})=-1.59$.

Figure 4

DGP3: Distributions Conditional on $x_{t-1}=.4$ and $w_{t-1}=.6$

$G_1(y_{t-1}) = 0$ $G_2(y_{t-1}) = 0.54$ $G_3(y_{t-1}) = 0.07$ $G_4(y_{t-1}) = 0.38$ $x^* = 0.6$ $w^* = -0.4$

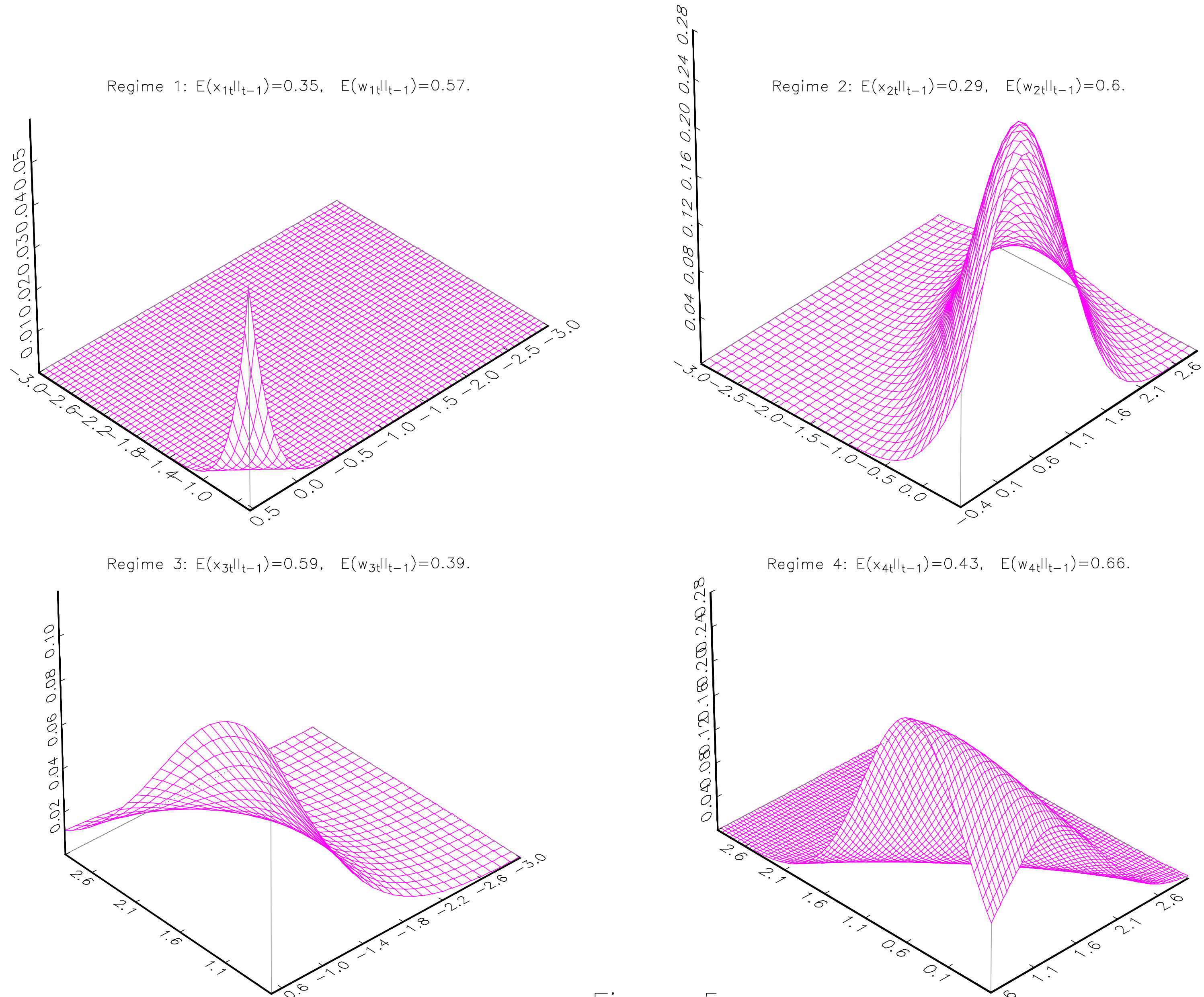

Figure 5

DGP3: Distributions Conditional on $x_{t-1}=-1.5$ and $w_{t-1}=-2$

$G_1(y_{t-1}) = 0.98$ $G_2(y_{t-1}) = 0.00$ $G_3(y_{t-1}) = 0.02$ $G_4(y_{t-1}) = 0$ $x^* = 0.6$ $w^* = -0.4$

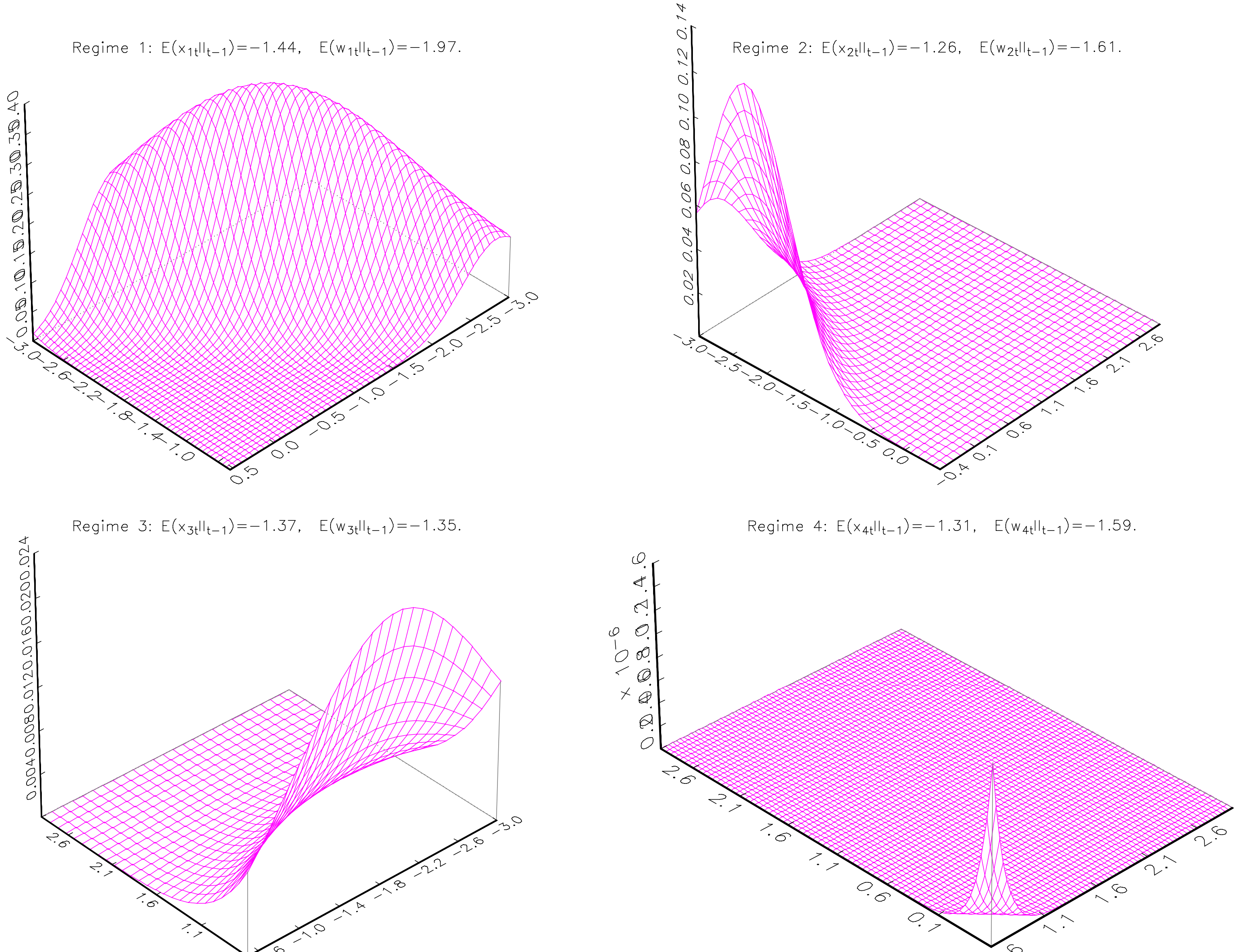

Figure 6

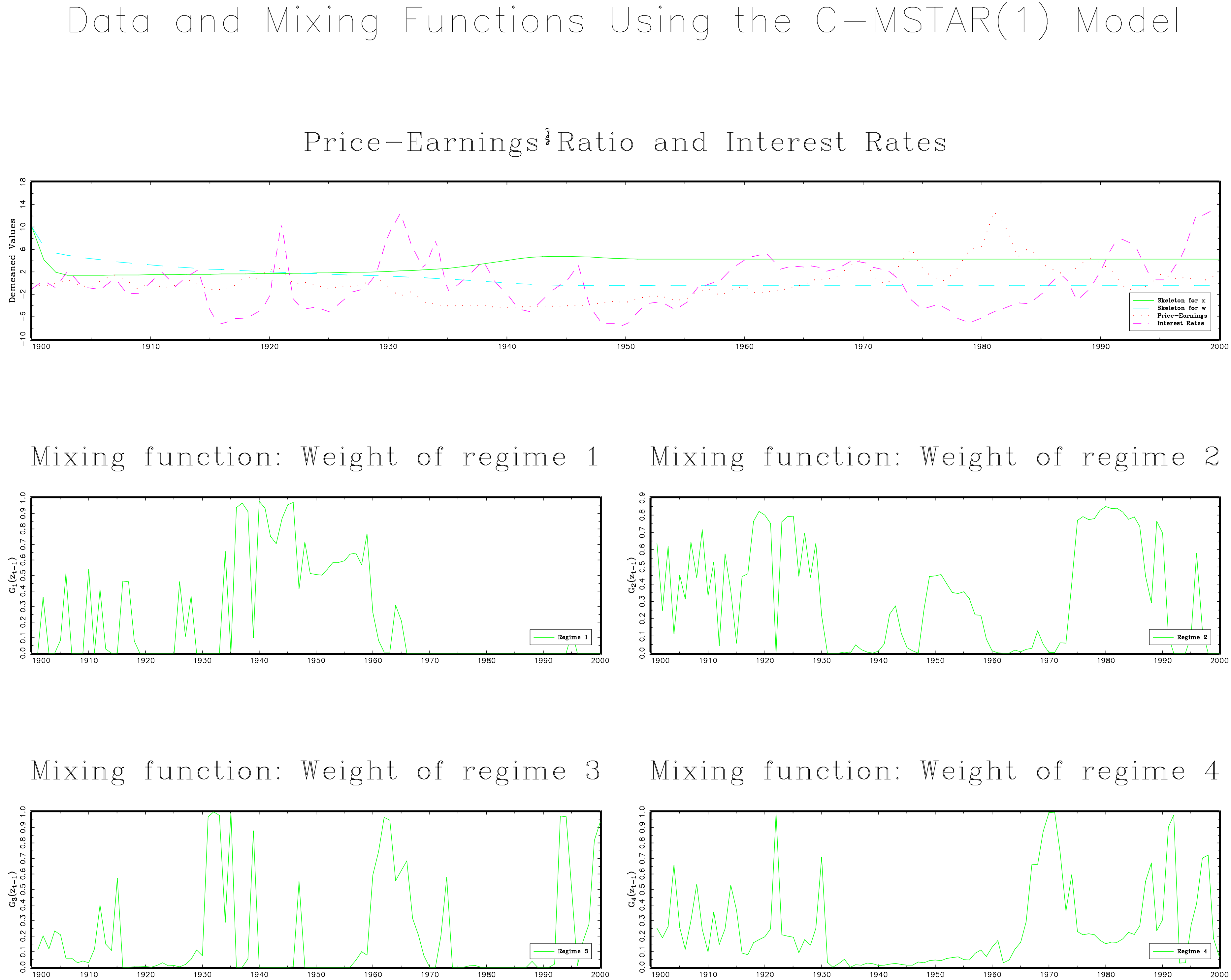
Data and Mixing Functions Using the C-MSTAR(1) Model
Price-Earnings Ratio and Interest Rates
Demeaned Values
Skeleton for x
Skeleton for w
Price-Earnings
Interest Rates
Mixing function: Weight of regime 1
Regime 1
Mixing function: Weight of regime 2
Regime 2
Mixing function: Weight of regime 3
Regime 3
Mixing function: Weight of regime 4
Regime 4

Figure 7